Salifort Motors Employee Turnover Analysis

Investigating the cause of turnover

OVERVIEW

Salifort Motors' leadership is concerned about the high turnover rate of employees. They asked the data team to analyze survey data and come up with ideas to increase employee retention. To do this, we built a machine learning model that can predict if an employee will leave based on the survey data. We chose to use an XGBoost Classifier to achieve this.

PROJECT STATUS

EDA was conducted on the survey data:
Data was cleaned: columns renamed
Duplicate data rows were kept, as they are most likely legit entries.

There were a significant number of outliers for tenure, however XGBoost Classifiers do not require eliminating outliers.

The survey data consists of 23.81% entries where the employee left the company.

An XGBoost Classifier was chosen as the model as it does not have a multicollinearity assumption like logistic regression.

After salary and department were converted into dummy variables, the XGBoost Classifier was built.

NEXT STEPS

Based on the fitted model, I would suggest further investigating how average monthly hours impacts employee satisfaction and their likelihood to quit.

Getter a better understanding of employee satisfaction, rather than a single 0-1 rating, may prove to be invaluable when trying to increase employee happiness and retention.

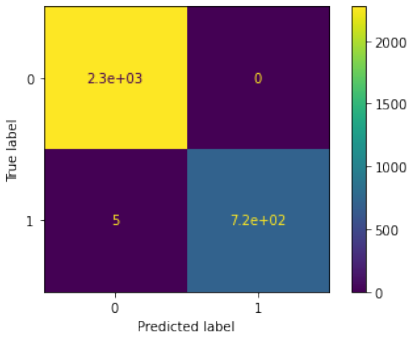
Gaining an understanding on how average monthly hours impacts employee satisfaction will also prove necessary for decision making.

I would not suggest making any decisions that may impact employees based on these results alone.

KEY INSIGHTS

The Fitted model has an accuracy of 99.83%. On a testing set with 3000 entries, it correctly classified all but 5.

All 5 errors were false negative errors. This results in an f1 score of 0.9965.



By observing the F-Score representing feature importance, we can conclude that average monthly hours has the most influence on making a correct prediction. Last evaluation score and the self reported satisfaction are a close second and third.

The department does not appear to be important in determining if an employee will quit.